Identification of diverse archaeal proteins with class III signal peptides cleaved by distinct archaeal prepilin peptidases

Szabó, Zalán; Oliveira Stahl, Adriana; Albers, Sonja-V.; Kissinger, Jessica C.; Driessen, Arnold J.M.; Pohlschröder, Mechthild; Pohlschroder, M.

Published in:
Journal of Bacteriology

DOI:
10.1128/JB.01547-06

Publication date:
2007

Table S2. Composition of operons containing FlaFind positive genes

Extent of gene locus	Number of genes	Number of substrates	*fla**	*tadA*	*tadC*	*flaK*	*eppA*	Other proteins encoded by operon/additional comments
Crenarchaeota								
***Aeropyrum pernix*K1**								
APE0697-APE0695	3	1		✓				
APE1907-APE1895	8	2	✓	✓	✓			
***Pyrobaculum aerophilum* str. IM2**								
PAE2394-PAE2414	16	2		✓				
***Sulfolobus acidocaldarius* DSM 639**								
Saci_1178-Saci_1172	7	1	✓	✓	✓			
Saci_1493-Saci_1500	9	2[†]		✓	✓			Predicted EndoIII-related nuclease (COG0177), Lhr-like helicase (COG1201), Glycosyltransferase probably involved in cell wall biogenesis (COG1215), Predicted transcriptional regulators (COG1475)
***Sulfolobus solfataricus* P2**								
SSO0122-SSO0117	6	2		✓	✓			tRNA and rRNA cytosine-C5-methylases (COG0144)
SSO2323-SSO3215	8	1	✓	✓	✓			operon disrupted by IS element integration
SSO2686-SSO2678	7	2		✓	✓			
SSO3142-SSO3138	5	3						
***Sulfolobus tokodaii* str. 7**								
ST1074-ST1070	5	3						
ST1396-ST1400	5	2		✓	✓			
ST2518-ST2524	7	1	✓	✓	✓			
Euryarchaeota								
***Archaeoglobus fulgidus* DSM 4304**								
AF0763-AF0758	6	4						
AF1055-AF1037	19	2	✓	✓	✓			Chemotaxis signal transduction system (AF1046-AF1037)
Haloarcula marismortui ATCC 43049								
rrnAC1252-rrnAC1244	9	5		✓	✓			Aspartate/tyrosine/aromatic aminotransferase (COG0436)
rrnAC1951-rrnAC1945	6	3		✓	✓			
rrnAC2203-2183	16	1	✓	✓	✓			Chemotaxis signal transduction system (rrnAC2194-

								rrnAC2192)
rrnAC1420-rrnAC1423	5	2						One gene located between rrnAC1422 and rrnAC1423 (similar to rrnAC1245) not annotated in published sequence
rrnAC1425-rrnAC1424	2	2						
***Halobacterium sp.* NRC-1**								
VNG0200C-VNG0198H	7	3						one gene located between VNG0204H and VNG0200C (similar to rrnAC1948) not annotated in published sequence
VNG0962G -VNG0953C	8	3[†]	✓					
VNG1008G-VNG1009G	2	2	✓					
***Methanocaldococcus jannaschii* DSM 2661**								
MJ0836-MJ0832.1	7	2					✓	Part of split operon, see Figure 3
MJ0891-MJ0906	16	3	✓	✓	✓	✓		
MJ1291 -MJ1282.1	12	1		✓	✓	✓		
MJ1467-MJ1472	7	2[†]						Part of split operon, see Figure 3
***Methanococcus maripaludis* S2**								
MMP0231-MMP0241	11	3					✓	See Figure 3
MMP0602-MMP0599	4	2						Orotidine-5'-phosphate decarboxylase (COG0284), Precorrin-6x reductase (COG2099)
MMP1666-MMP1676	11	3	✓	✓	✓			
***Methanopyrus kandleri* AV19**								
MK0704-MK0698	7	2					✓	upstream genes not directly linked to this locus encode chaperone-like ATPase (COG0606), Flp pilus assembly ATPase (COG4962) and two TadC homologs (COG2064)
MK0696-MK0694	3	1						gene cluster possibly co-regulated with above listed operon
***Methanosarcina acetivorans* C2A**								
MA1234-MA1230	5	3						ABC transporter components
MA3062-MA3054	9	2	✓	✓	✓			
MA3077-MA3084	8	1	✓	✓	✓			second *fla* operon similar to MM0418-MM0411
MA3093-MA3098	6	1		✓	✓			ABC transporter components
***Methanosarcina mazei* Go1**								

MM0323-MM0315	9	2	✓	✓	✓			upstream gene locus codes for chemotaxis system
MM0418-MM0411	8	1	✓	✓	✓			second *fla* operon similar to MA3077-MA3084
Methanothermobacter thermautotrophicus str. Delta H								
MTH421-MTH432	12	4[†]					✓	conserved membrane protein (COG4025), predicted PP loop family ATPase (COG2102)
***Pyrococcus abyssi* GE5**								
PAB1376-PAB1388	13	3	✓	✓	✓			SAM dependent methyltransferase (COG0500), conserved archaeal protein (COG3373)
PAB1817-PAB2392	12	6						(Lhr-like helicase (COG1201), cell division GTPase COG0206), uncharacterized conserved protein COG5306), uncharacterized membrane protein (COG4505)
***Pyrococcus furiosus* DSM 3638**								
PF0340-PF0329	12	2	✓	✓	✓			uncharacterized archaeal protein (COG3373), SAM dependent methyltransferase (COG0500)
PF0472-PF0469	4	1					✓	uncharacterized conserved proteins (COG4043/COG2411)
PF1307-PF1303	5	2						
PF1504-PF1511	8	2						Lhr-like helicase (COG1201), cell division GTPase (COG0206)
***Pyrococcus horikoshii* OT3**								
PH0544-PH0560	14	5	✓	✓	✓			uncharacterized archaeal protein (COG3373), SAM dependent methyltransferase (COG0500)
PH1329-PH1344	12	5						Lhr-like helicase (COG1201), cell division GTPase (COG0206), uncharacterized protein (COG5306), uncharacterized membrane protein (COG4505)
***Thermococcus kodakaraensis* KOD1**								
TK0036-TK0050	13	5	✓	✓	✓			
TK1670-TK1661	10	5						Lhr-like helicase (COG1201), cell division GTPase (COG0206)
TK1707-TK1695	11	2						
***Thermoplasma acidophilum* DSM 1728**								
Ta0553-Ta0556	4	1	✓					
***Thermoplasma volcanium* GSS1**								
TVN0607-TVN0614	8	1	✓	✓	✓			

***Nanoarchaeum equitans* Kin4-M**								
NEQ137-NEQ132	5	2						

*flagellar accessory genes

†substrate(s) in these operons were included after reannotation: VNG0961G, VNG0962G and VNG1009G (missing leader), MJ1467 (long leader), MMP1666 (missing leader), MTH422 (long leader), PF0338 (missing leader), Saci_1496.1 (not annotated as gene)

Table S3. Plasmids used in this study

Plasmid	Description	Source
pBAD/*Myc*-His A	*E. coli* expression vector, arabinose inducible promoter	Invitrogen
pZA7	T7 expression vector, C-terminal HA tag	Ref x
pSA4	T7 expression vector, C-terminal 6xHIS tag	Ref x
pZA10	pZA7 with *epdA*	This study
pZA11	pZA7 with *epdB*	This study
pZA12	pZA7 with *flaB2*	This study
pZA13	pSA4 with *eppA*	This study
pZA14	pSA4 with *flaK*	This study
pZA10-*eppA*	*epdA*/*eppA* co-expression	This study
pZA10-*flaK*	*epdA*/*flaK* co-expression	This study
pZA10-*pibD*	*epdA*/*pibD* co-expression	This study
pZA11-*eppA*	*epdB*/*eppA* co-expression	This study
pZA11-*flaK*	*epdB*/*flaK* co-expression	This study
pZA11-*pibD*	*epdB*/*pibD* co-expression	This study
pZA12-*eppA*	*flaB2*/*eppA* co-expression	This study
pZA12-*flaK*	*flaB2*/*flaK* co-expression	This study
pZA12-*pibD*	*flaB2*/*pibD* co-expression	This study
pZA13	pZA7-*epdA*(RGQI → KGAS)	This study
pZA14	pZA7-*epdB*(KGQV → KGAS)	This study
pZA15	pZA7-*flaB2*(KGAS → RGQI)	This study
pZA13-*eppA*	$epdA^{KGAS}$/*eppA* co-expression	This study
pZA13-*flaK*	$epdA^{KGAS}$/*flaK* co-expression	This study
pZA14-*eppA*	$epdB^{KGAS}$/*eppA* co-expression	This study
pZA14-*flaK*	$epdB^{KGAS}$/*flaK* co-expression	This study
pZA15-*eppA*	$flaB^{RGQI}$/*eppA* co-expression	This study
pZA15-*flaK*	$flaB^{RGQI}$/*flaK* co-expression	This study
pZA16	pZA7-*edpA*-6xHIS	This study
pZA17	pZA7-*edpB*-6xHIS	This study
pZA18	pZA7-*flaB2*-6xHIS	This study
pZA16-*eppA*	*edpA*-6xHIS/*eppA* co-expression	This study
pZA17-*eppA*	*edpB*-6xHIS/*eppA* co-expression	This study
pZA18-*flaK*	*flaB2*-6xHIS/*flaK* co-expression	This study